

Electronic Supplementary Material

Biodiversity conservation in agriculture requires a multi-scale approach

David J. Gonthier, Katherine K. Ennis, Serge Farinas, Hsun-Yi Hsieh, Aaron L. Iverson, Péter Batáry, Jörgen Rudolphi, Teja Tscharntke, Bradley J. Cardinale, Ivette Perfecto

S1 Text

Selection of studies

The keyword search in ISI Web of Science consisted of the following string of keywords: (biodiversity OR diversity OR abundance*) AND (landscape OR regional) AND agr* AND scale* AND (local OR management OR intensity) AND (bird* OR avian* OR arthropod* OR insect* OR weed* OR bat* OR mammal* OR butterfly* OR rodent* OR reptile* OR herp* OR plant* OR spider* OR invert* OR pollinator* OR fung* OR parasite* OR “seed dispers*” OR “biological control” OR “natural enem*” OR predator* OR parasitoid* OR pest*)).

Publication bias

Failure to publish negative or non-significant results with low samples size can result in literature for which outcomes are bias and strongly positive. Therefore in quantitative syntheses it has become commonplace to test for the importance of publication bias using a number of methods. If a correlation between sample size and effect size exists many argue this is evidence for bias toward publication of studies with positive effects with large sample sizes. Failsafe numbers are also used to estimate the number of studies with null results needed to eliminate the significance of a statistical analysis [1]. Across our observations there were no significant correlations between effect size and n for management LR_M or landscape Z_L for richness and abundance data (Table S1). Rosenthal’s fail-safe analysis suggested that at least 100 nil observations were needed to eliminate statistical significance ($p < 0.05$) across all analyses with significant mean effect sizes, except in one case. The fail-safe number for the landscape Z_L of vertebrate richness was only 32.5, therefore we caution the interpretation of this result, but note that 32.5 nil observations are still 3.6 times more observations than the number of existing observations.

Table S1. Correlations between sample size and effect size and Rosenthal's fail-safe numbers.

	Model $P < 0.05$	Effect size vs. n n	R	p	Fail safe x^\dagger	x/n
Richness						
Local Management LR_M						
Overall	Yes	70	-0.130	0.284	3220.3	46.0
Plant	Yes	13	-0.157	0.610	703.2	54.1
Invertebrate	Yes	46	-0.108	0.466	627.7	13.6
Vertebrate	No	9	-0.096	0.806	12.2	1.4
Landscape Z_L						
Overall	Yes	71	-0.052	0.668	1256.9	17.7
Plant	No	14	-0.011	0.970	17.9	1.3
Invertebrate	Yes	46	-0.036	0.808	544.9	11.8
Vertebrate	Yes	9	-0.420	0.260	32.5	3.6
Abundance						
Local Management LR_M						
Overall	Yes	62	-0.226	0.076	1504.9	24.3
Plant	No	8	-0.395	0.333	0.0	0.0
Invertebrate	Yes	46	-0.196	0.191	748.8	16.3
Vertebrate	No	9	-0.116	0.767	97.0	10.8
Landscape Z_L						
Overall	Yes	63	-0.210	0.098	148.4	2.4
Plant	No	8	0.512	0.194	0.0	0.0
Invertebrate	Yes	46	-0.276	0.063	109.1	2.4
Vertebrate	No	9	-0.268	0.485	0.0	0.0

† Fail safe number (x) was calculated from effect size and variance as in [1].

Table S2. Statistical tests of fixed effects for richness and abundance models weighted by the inverse of variance. For all models a random effect of study was included. Statistical models were used to estimate mean and 95% confidence interval of effect sizes for overall responses and taxonomic groups (Table S3).

	<i>d.f.*</i>	<i>F</i>	<i>P</i> [†]
Local Management LR_m^{\dagger} Richness			
Intercept	1,39	19.3	<0.001
Taxonomic group	1,54	4.9	0.012
Crop type	3,29	1.1	0.351
Local Management LR_m^{\dagger} Abundance			
Intercept	1,23	6.3	0.019
Taxonomic group	2,29	0.7	0.487
Crop type	3,19	2.2	0.116
Landscape Z_L^{\ddagger} Richness			
Intercept	1,33	9.8	0.004
Taxonomic group	2,54	0.3	0.737
Crop type	3,36	2.4	0.084
Landscape Z_L^{\S} Abundance			
Intercept	1,24	8.2	0.009
Taxonomic group	2,34	<0.1	0.997
Crop type	3,14	2.1	0.154
Landscape factor type	1,49	3.8	0.056

**d.f.* = *numerator, denominator*. [†]For two studies, two time points were recorded for richness or abundance across year or season, we therefore summed means and standard errors before calculating local management LR_M and estimated variance in LR_M from these summed standard errors and assumed a sample size equivalent to the number of sites. For landscape factor Z_L we averaged across the two time points and estimated the variance in Z_L assuming a sample size equivalent to the number of sites. Inclusion or exclusion of these studies had no influence on the statistical outcomes; therefore we did not consider more complicated estimations of effect size variances.

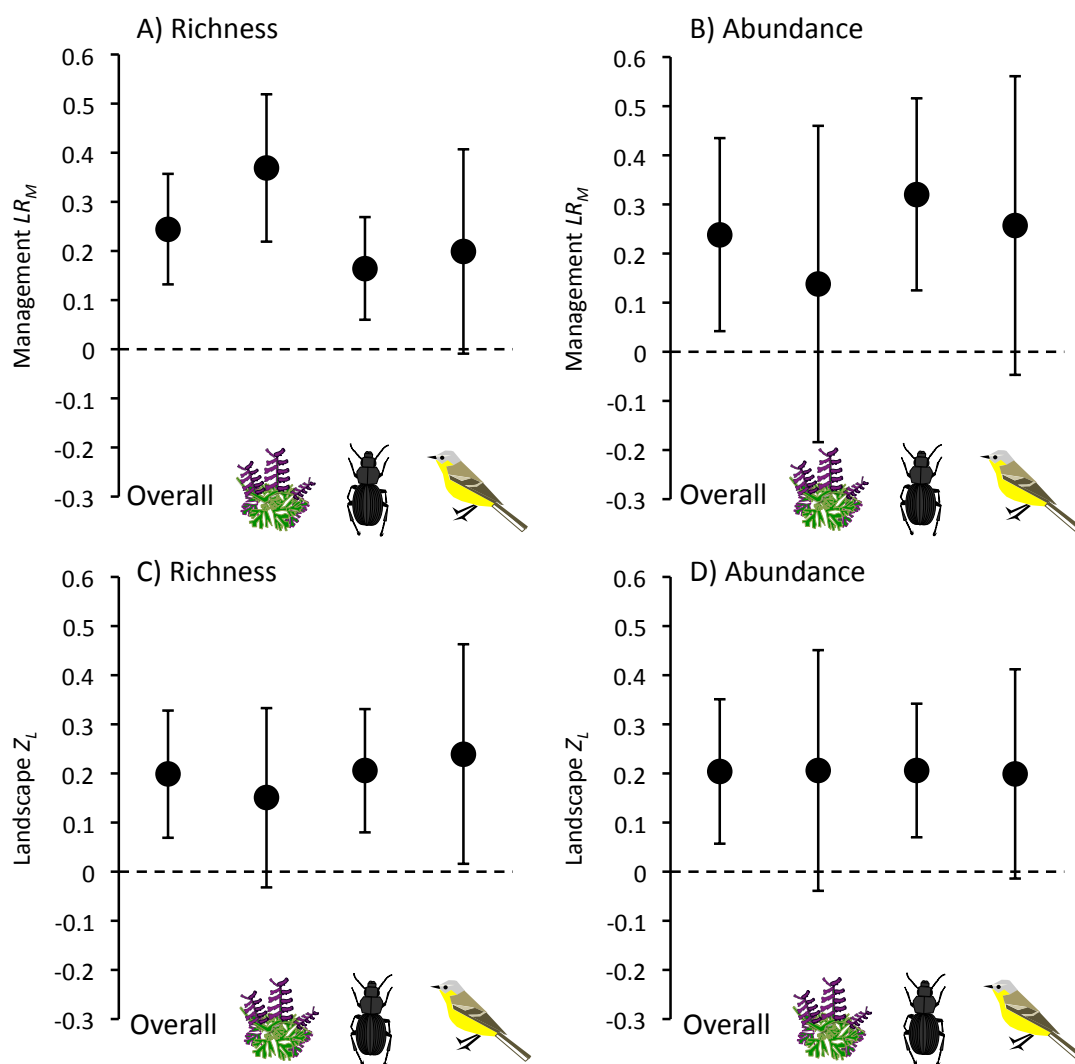


Figure S1. Estimated marginal means and confidence intervals (95%-CI) for models of richness and abundance weighted by the inverse of variance. Summary statistics of the GLMMs used to estimate marginal means and 95% CIs are available in Table S2.

Figure S2

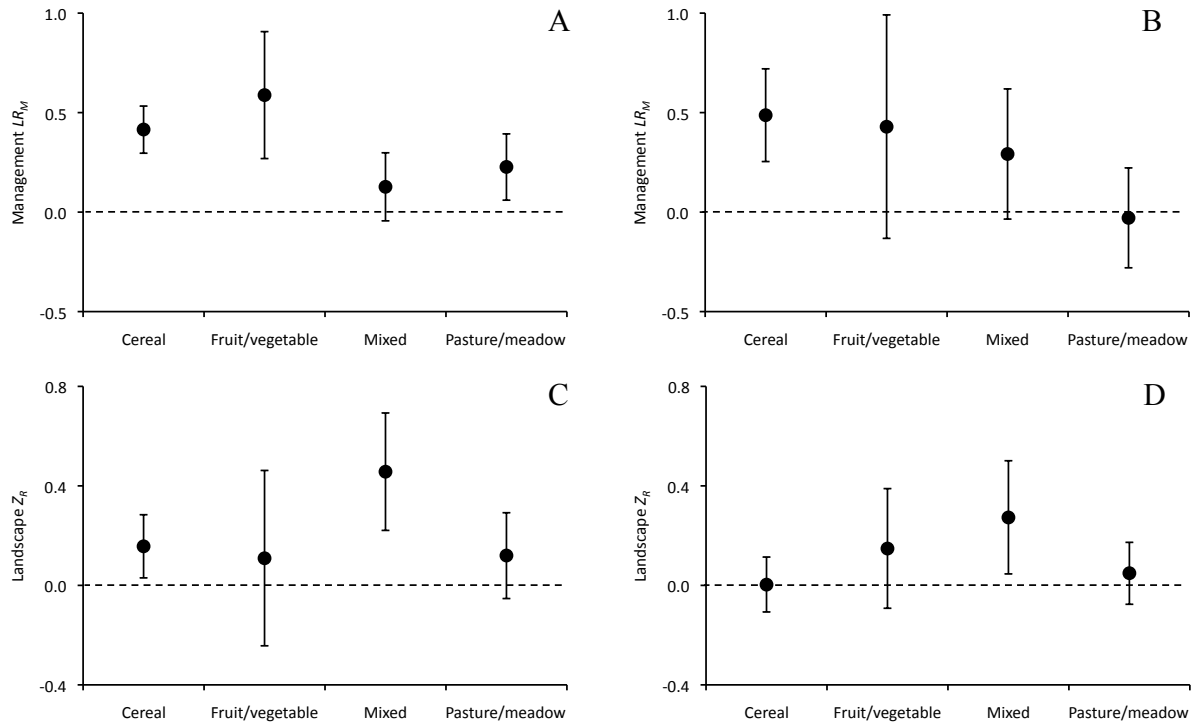


Figure S2. Management and landscape effects on richness and total abundance across type of agriculture. Estimated marginal mean and 95% CI for agricultural types: cereal, fruit/vegetable, mixed, and pasture/meadows for management LR_M for richness (A) and abundance (B) and for landscape Z_L for richness (C) and abundance (D).

References:

1. Rosenthal R. 1979 The "file drawer problem" and tolerance for null results. *Psychological Bulletin* **86**, 638-641.